

Credo: Declarative Control of LLM Pipelines via Beliefs and Policies

Duo Lu (Brown) Andrew Crotty (Northwestern) Uğur Çetintemel (Brown)

GitHub Repo: <https://github.com/brown-db/credo>

- What?**
- A semantic control plane for LLM pipelines, stored and inspectable as structured data.
- Why?**
- A fixed plan must either overprovision for easier tasks or risk failure on harder ones.
- How?**
- Beliefs* are semantic facts about state; declarative *policies* refine the plan based on them.

1 The Problem: One Size Does Not Fit All for LLM Pipelines

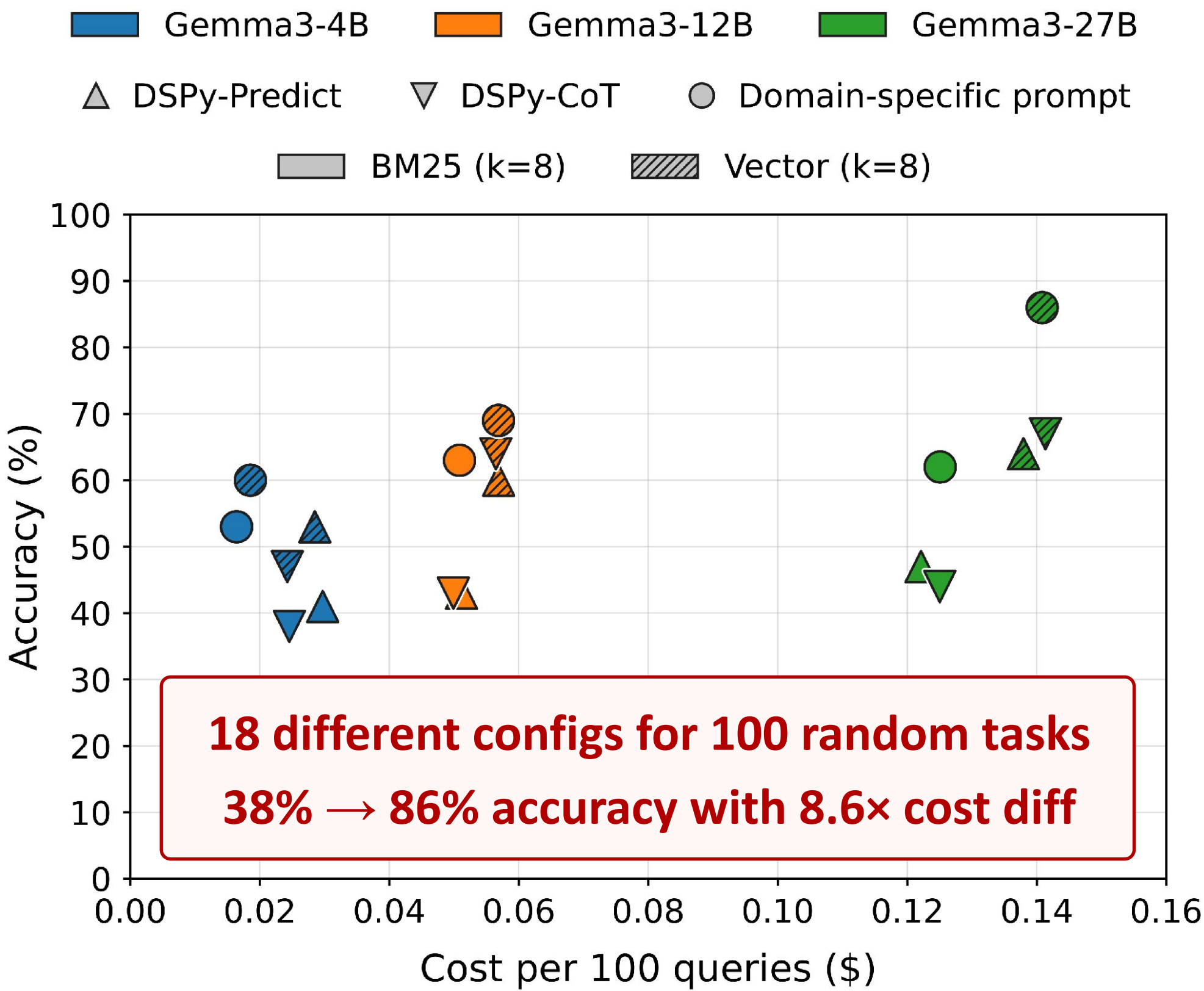
Same pipeline for different questions in a workload.

- “What was Apple’s total revenue in FY2022?”**
simple lookup/number extraction + a small model (SLM) is sufficient
- “What is Microsoft’s FY2022 gross profit margin?”**
extracting values from different tables + arithmetic requires LLM

Existing approaches in this space fall short.

- Declarative but static**
LOTUS · DocETL · Palimpzest
plan frozen at compile time
- Adaptive but opaque**
ReAct · AutoGen
logic buried in prompts

Our main idea: observe, decide, and adapt at runtime.



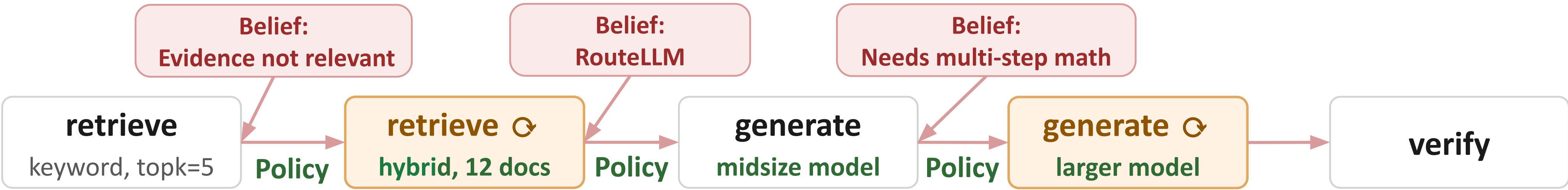
2 Pipeline Walkthrough – “What is Microsoft’s FY2022 gross profit margin?”

Without Credo (a vaillina LLM pipeline — one fixed plan for every question)

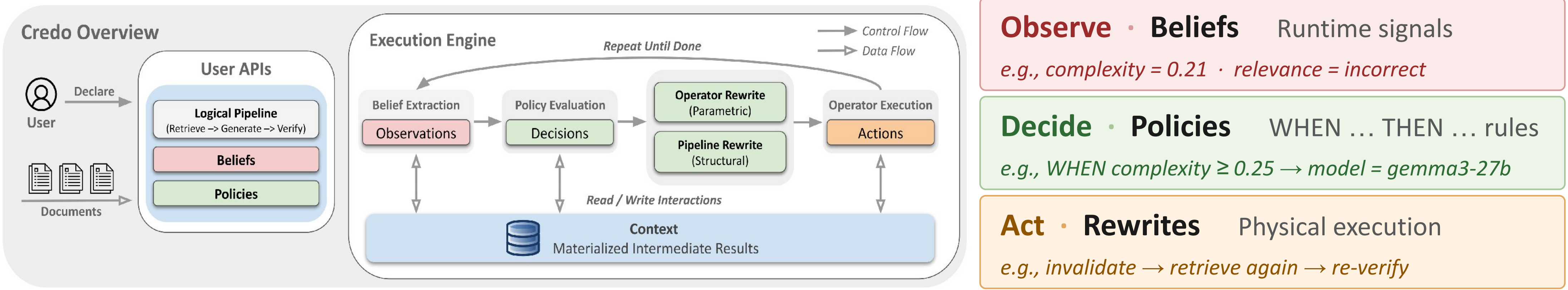


Make the plan adaptive at runtime based on user-defined semantic signals (*beliefs*) and declarative rules (*policies*).

With Credo (*beliefs* capture runtime signals and *policies* trigger adaptive rewrites)

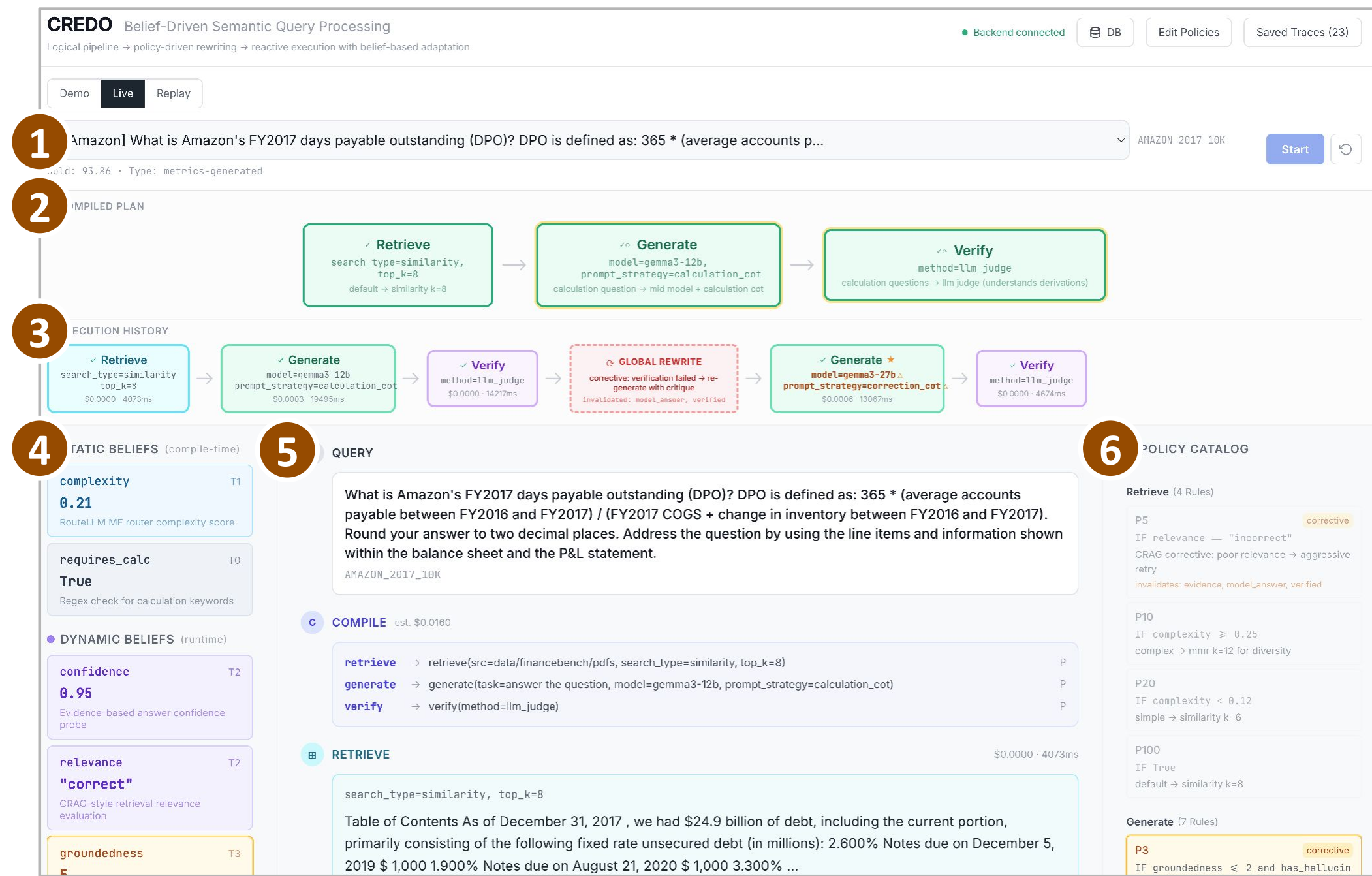


3 Credo: A Semantic Control Plane for LLM Pipelines



- Observe · Beliefs** Runtime signals
e.g., complexity = 0.21 · relevance = incorrect
- Decide · Policies** WHEN ... THEN ... rules
e.g., WHEN complexity ≥ 0.25 → model = gemma3-27b
- Act · Rewrites** Physical execution
e.g., invalidate → retrieve again → re-verify

4 Overview of the Demo Web Interface



- 1 Task selector
- 2 Generated logical plan
- 3 Physical execution history
- 4 Belief catalog
- 5 Detailed step trace
- 6 Policy catalog



Paper



GitHub Repo